

# Automatic in Vivo Cell Detection in MRI

Muhammad Jamal Afridi<sup>1</sup>, Xiaoming Liu<sup>1</sup>, Erik Shapiro<sup>2</sup>, and Arun Ross<sup>1</sup>

<sup>1</sup> Department of Computer Science and Engineering,  
Michigan State University

<sup>2</sup> Department of Radiology, Michigan State University  
{afridimu, liuxm, shapir86, rossarun}@msu.edu

**Abstract.** Due to recent advances in cell-based therapies, non-invasive monitoring of *in vivo* cells in MRI is gaining enormous interest. However, to date, the monitoring and analysis process is conducted manually and is extremely tedious, especially in the clinical arena. Therefore, this paper proposes a novel computer vision-based learning approach that creates superpixel-based 3D models for candidate spots in MRI, extracts a novel set of superferm features, and utilizes a partition-based Bayesian classifier ensemble to distinguish spots from non-spots. Unlike traditional ferns that utilize pixel-based differences, superferns exploit superpixel averages in computing difference-based features despite the absence of any order in superpixel arrangement. To evaluate the proposed approach, we develop the first labeled database with a total of more than 16 thousand labels on five *in vivo* and four *in vitro* MRI scans. Experimental results show the superiority of our approach in comparison to the two most relevant baselines. To the best of our knowledge, this is the first study to utilize a learning-based methodology for *in vivo* cell detection in MRI.

## 1 Introduction

As a promising alternative to organ transplants in humans, cell transplant-based therapies have recently gained enormous interest in medical research. However, its long term success in humans has not been proven, where one key obstacle is the ability to non-invasively monitor the transplanted cells via MRI, and measure both transplant efficiency and long-term cell repopulation. Overcoming this obstacle is equivalent to achieving two goals: (i) display the transplanted cells in MRI scans under *in vivo* conditions, and (ii) perform comprehensive quantitative analysis on the behavior of transplanted cells. To display transplanted cells in MRI scans, cells are injected with an MRI contrast agent prior to or during imaging [8]. This has been demonstrated in many *in vivo* studies, e.g., [7,2].

While such cell studies use efficient magnetic particles and high-resolution MRI, they stopped short of quantification and hence cannot accomplish the second goal. For comprehensive quantitative analysis, *manual* enumeration of cells in MRI is *laborious*, *subjective*, and *limited* in capturing all patterns of cell behaviors. Further, current commercial tools can only assist humans in conducting

manual analysis. Therefore, there is a need to automatically and accurately perform such quantitative analysis of cells or spots. Note that we use *cell* or *spot* interchangeably since a cell visually appears as a spot in MRI scans.

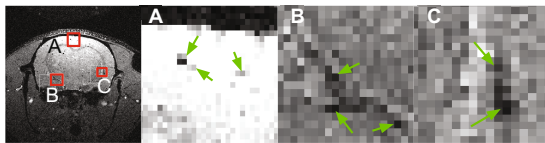
Recognizing this need, a recent study utilized a simple threshold-based strategy to automatically detect cell spots in 2D MRI images of a rat’s brain [4]. However, they did not quantitatively report the spot detection accuracy against any ground truth. In contrast, supervised machine learning-based solutions are more robust and have recently shown enormous success in a wide range of medical applications. Unfortunately, there is no prior work that utilizes learning-based approaches for automatically detecting *in vivo* spots in MRI. In a related area of 2D microscopic image analysis, authors in [9] conducted a comparative study of various machine learning approaches in detecting spots. The study concluded that a Haar-based Adaboost approach performs the best on their data. Nevertheless, *in vivo* 3D MRI scans present more challenging in nature than fluorescence based microscopic images of [9]. E.g., MRI scans also contain spot-like tissue’s structural entities in the image background that are absent in microscopic images. Secondly, based on the size and shape of the spots in [9], HAAR-like features are extracted from a fixed sized mask that slides on a 2D image, which may not work for our 3D spots due to the large size and shape variations of spots.

Moving towards accomplishing the second goal, this paper presents the first comprehensive study on learning-based detection of cell spots in MRI scans. Our cell detection framework has novelty in terms of spot modeling, feature representation, and classification. Our approach considers spots as 3D entities and represents its general structural model using superpixels. We then extract a novel set of “superferns” features and finally classify it using a partition-based ensemble of Bayesian networks. Experimental results show that our proposed approach performs significantly better than previously related approaches.

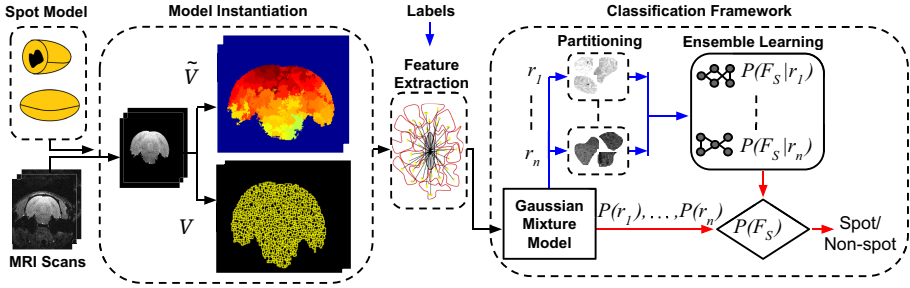
In summary, this paper makes the following contributions: (i) It proposes a novel superpixel-based 3D model to characterize cellular spots and that can potentially be used in other medical problems. (ii) It introduces the *superferns* feature that exploits superpixel-based representations and is more discriminative than traditional fern features. (iii) It demonstrates *how* a partition-based ensemble learning can be effectively utilized for MRI spot detection. (iv) It presents a labeled publicly available database of five *in vivo* and four *in vitro* MRI scans, and a total of 16, 298 spot labels for quantitative evaluation.

## 2 Proposed Approach

The cell/spot detection problem in MRI scans has unique challenges, where a number of questions should be carefully considered prior to algorithm design. First, since a spot is essentially a 3D entity in an MRI cube, *how* can we model



**Fig. 1.** Spot variations in diverse region contexts.



**Fig. 2.** Our approach has four main modules. Blue, red, and black arrows are the processing flow during the training stage, testing stage and both stages, respectively.

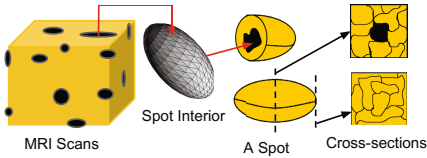
its three dimensional characteristics? Second, a spot is also a small group of dark pixels with varying shapes and sizes. *What is the basic unit* within an MRI cube (e.g., one, two, or  $N$  pixels) for which the two-class classification decision can be made? Third, there is a huge number of candidate locations. So, *how* to construct discriminative and efficient feature representation for spots? Fourth, the appearance of a spot varies relative to its local and regional neighborhood (Fig. 1). *How* to make learning robust to these variations should be addressed. Considering these challenges, we design our technical approach as in Fig. 2.

## 2.1 Spot Modeling

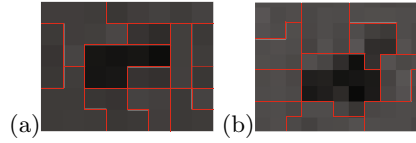
Visually, a cellular spot  $\mathbf{S}$  appears as a cluster of dark 3D pixels with high variations in its 3D shape and intensity, wrapped inside a cover of background pixels as conceptually illustrated in Fig. 3. In this work, we call the small group of dark pixels as a spot’s interior  $I$ , and their local neighboring pixels in the background as the exterior  $E$  of a 3D spot. This model is consistent with the manual labeling of spots by domain experts, who inspect the cross-sections of spots in consecutive 2D MRI slices, and look for a small region (interior) that is darker than its neighboring pixels (exterior). Furthermore, the human visual system can also adjust for the amount of relative darkness based on the characteristics of the *specific brain region* containing that spot. E.g., Fig. 1 shows that the way to classify spots in region A and B might be different, and one spot in C is comparatively larger. Thus, in addition to modeling a spot with its interior/exterior, we also model the specific region it belongs to, termed *region context*  $R$ .

## 2.2 Model Instantiation via Superpixel

Given the conceptual spot model  $\mathbf{S} = \{I, E, R\}$ , we now describe how to define  $I$ ,  $E$ , and  $R$  for a spot, by three steps. Since no spot should be outside the brain region, the first step is to perform brain segmentation in every 2D MRI slice with basic image processing techniques. The second step is to define  $I$  and  $E$  by applying 2D superpixel extraction [3] to the segmented brain



**Fig. 3.** Conceptual Spot Model



**Fig. 4.** Superpixels capture a tight (a) or loose (b) boundary of a spot's interior.

region of each slice. A superpixel is a group of  $N$  neighboring pixels with similar intensities, i.e.,  $V_{z,u} = \{x_i, y_i, z\}_{i=1}^N$  where  $u$  is the superpixel ID in slice  $z$ . In general superpixels can tightly capture the boundary of a spot's interior; however, some imprecise localization is also expected in practice (Fig. 4). After extraction, we denote  $\mathbb{M} = \{V_{z,u}\}_{z=u=1}^{L,U}$  as the set of all superpixels in the 3D brain region, where  $L$  is the number of slices and  $U$  is the number of superpixels per slice. Due to the exclusiveness of the interior and exterior of spots, we have  $\mathbb{M} = \mathbb{I} \cup \mathbb{E}$  where  $\mathbb{I}$  and  $\mathbb{E}$  are the set of all interior and exterior superpixels, respectively. With that, for a spot  $\mathbf{S}$  with length  $l$  in  $z$ -axis, we formally define its interior as  $I = \{V_{z,u}, \dots, V_{z+l-1,u} \mid V \subset \mathbb{I}\}$  and the exterior as  $E = \{V_{z-1,\bar{u}}, V_{z,\bar{u}}, \dots, V_{z+l-1,\bar{u}}, V_{z+l,\bar{u}} \mid \|(m(I) - m(V))\| \leq \tau, V \subset \mathbb{E}\}$ , where  $m()$  is the mean of a set,  $\tau$  is the maximum  $L^2$  distance between the centers of a spot and an exterior superpixel, and  $V_{z-1,\bar{u}}$  and  $V_{z+l,\bar{u}}$  are superpixels in two adjacent neighboring slices. To define the region context  $R$ , in the third step we extract very large superpixels  $\{\tilde{V}_{z,u}\}_{z=u=1}^{L,\tilde{U}}$ , where the number of superpixels  $\tilde{U} \ll U$ , by assuming that large superpixels are representative of the regional appearance. Thus, we define the region context of a spot as  $R = \{\tilde{V}_{z,u} \mid m(I) \subset \tilde{V}_{z,u}\}$ , which is the large superpixel enclosing the spot center  $m(I)$ . Consequently, this process allows each superpixel to act as an interior of a potential spot and therefore naturally generates a large number of candidates.

The superpixel-based 3D spot model has a few advantages. First, it addresses the issue of *unit*, by going beyond pixels and using superpixel-based feature extraction and classification. Second, it substantially limits the total candidate spots to be tested, since the candidates are nominated via superpixels rather than pixels. Note that we may extend our 3D spot model by directly using 3D supervoxels instead of joining 2D superpixels. We choose the latter in this work due to its superior reliability than well-known 3D supervoxel methods.

### 2.3 Superferns Feature Extraction

With an instantiated spot model  $\mathbf{S} = \{I, E, R\}$ , the next step is to extract a discriminative and efficient feature representation. Since a spot generally has darker interior than its exterior, it makes sense to define features based on the computationally efficient intensity differences between pixels in the interior and exterior. Difference-based fern features have shown great success in computer vision [5]. Ferns compute the intensity difference between a *subject pixel* and

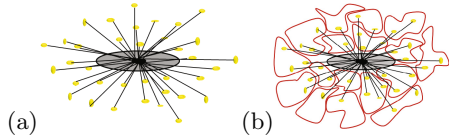
another pixel with a certain offset w.r.t. the subject pixel. Using the same offset in different images leads to feature correspondence among these images.

For our problem, the spot center  $m(I)$  can be regarded as the subject pixel, and its intensity is the average intensity of all interior pixels  $m(\mathbf{G}(I))$ . We then randomly generate  $h$  3D offsets  $O = \{\mathbf{o}_i\}_{i=1}^h$  with a uniform distribution, whose center is the spot center and radius is  $\tau$ . Finally, the feature set is computed as  $F = \{f_i\}_{i=1}^h$ , where  $f_i = \mathbf{G}(m(I) + \mathbf{o}_i) - m(\mathbf{G}(I))$ . While  $f_i$  is efficient to compute,  $\mathbf{G}(m(I) + \mathbf{o}_i)$  is the intensity of a single pixel, which can be noisy, specially in *in vivo* MRI and lead to low discriminability of  $f_i$ . Thus, it is desirable to replace it with the average intensity of all pixels within an exterior superpixel. However, the exterior superpixels around different spots have no correspondence, and, as a result,  $f_i$  for different spots also have the correspondence issue.

To address this issue, we present an approach to exploit the average intensity without losing correspondence information. The new feature, termed as ‘‘superferns’’, is similar to  $F$  except it replaces the single pixel-based intensity with the average intensity of the superpixel (Fig. 5 (b)),

i.e.,  $F' = \{f'_i\}_{i=1}^h$ , where  $f'_i = m(\mathbf{G}(V)) - m(\mathbf{G}(I))$ ,  $\forall m(I) + \mathbf{o}_i \in V$ . Note that it is possible to have the same feature at two different offsets due to them being in the same superpixel, i.e.,  $f'_i = f'_j$ . This is not an issue because this equality may not be true for other spots, hence the feature distributions of  $f'_i$  and  $f'_j$  are not the same, and they contribute differently to the classification.

Features are also needed for the region context  $R$ . Given its role of supporting region-dependent classifiers, we find that simple features work well for  $R$ , e.g., the mean and standard deviation of pixel intensities in  $R$ ,  $F_r = (m(\mathbf{G}(R)), \sigma(\mathbf{G}(R)))$ .



**Fig. 5.** The ferns (a) and superferns (b).

## 2.4 Partition-Based Bayesian Classification

Having computed the feature  $F_s = (F', F_r)$  for a set of spots and non-spots, we now present our approach to learn an accurate two-class classifier. Since different local regions have different appearance as shown in Fig. 1, we partition the brain region into  $N_0$  partitions, learn a set of  $N_0$  classifiers each for one partition, and fuse them via a probabilistic Bayesian formulation. Specifically, for any spot candidate  $\mathbf{S}$ , its probability of being a spot is

$$P(F_s) = \sum_{i=1}^{N_0} P(F_s, r_i) = \sum_{i=1}^{N_0} P(F_s|r_i)P(r_i), \quad (1)$$

where  $r_i$  represents the  $i^{\text{th}}$  partition,  $P(r_i)$  is the probability of  $\mathbf{S}$  belonging to  $r_i$ , and  $P(F_s|r_i)$  is the conditional probability of being a spot at partition  $r_i$ .

We learn  $P(r_i)$  using Gaussian Mixture Models (GMM) with random initialization. By collecting the  $F_r$  for all training samples, we perform GMM to estimate  $N_0$  component Gaussian densities (using Aikake Information Criterion),

each considered as one partition. During the testing,  $\{P(r_i)\}_{i=1}^{N_0}$  is obtained by evaluating  $F_r$  of the testing sample w.r.t. each component densities.

In order to learn  $P(F_s|r_i)$ , we group all training samples into  $N_0$  groups based on their respective maximum  $\{P(r_i)\}$ , and train the  $P(F_s|r_i)$  using the standard implementation of Bayesian Networks in [1], where the maximum number of parent nodes is set to 4. During the test, for a testing candidate spot, GMM enables a soft partition assignment, and its final probability of being a spot is the weighted average from the classifier ensemble.

### 3 MRI Database for Cell Detection

Since there is no public database for cell detection in MRI, we collect a database in both *in vivo* and *in vitro* settings, and manually label the ground truth spots.

***In vitro* MRI:** A phantom was constructed consisting of a known number of 4.5 micron diameter, magnetic microparticles with 10 pg iron per particle, suspended in 1% agarose. T2\*-weighted gradient echo MRI was performed on this sample at a resolution of 100 micron isotropic voxels ( $128 \times 80 \times 80$  pixels). Using a field strength of 7T, a total of four MRI *in vitro* scans of different tubes were developed. One of these tubes was naïve whereas the rest contained spots.

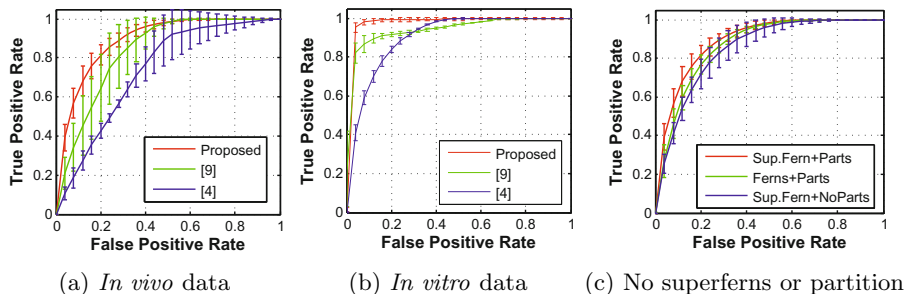
***In vivo* MRI:** At a field strength of 11.7T, five rat brains are scanned to obtain MRI cubes of  $256 \times 256 \times 256$  pixels. Three rats were injected intracardiac with MPIO-labeled MSCs, delivering cells to the brain, and two rats were naïve.

**Labeling:** For quantitative evaluation, for every MRI scan with spots, a domain expert manually clicked on a set of 2D dark pixels on every slice, using our GUI labeling tool that allows zooming in locally. The superpixel of one clicked pixel, or the concatenation of multiple nearby pixels, forms the interior of one ground truth 3D spot. This leads to a set of ground truth spots  $S = \{\mathbf{s}_i\}_{i=1}^K$ ,  $\mathbf{s}_i = (x_i, y_i, z_i)^T$ , where  $x_i, y_i, z_i$  are the 3D spot center. There are 841 and 15,457 ground truth spots on three *in vitro* scans and three *in vivo* scans, respectively. Labeling such data is not only time consuming but also *subjective*. Therefore, to learn the inter-rater variability, we compute  $Dv = 1 - \frac{1}{3} \sum_{i=1}^3 |J1 \cap J2| / |J1 \cup J2|$  to be  $\sim 36\%$  on three *in vitro* scans, where  $J1$  and  $J2$  are labels of two experts.

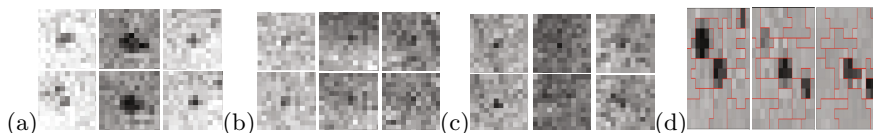
### 4 Experimental Results

In this section we design experiments to investigate answers to the following questions: (i) how does our approach perform and compare with the previous approaches using both *in vivo* and *in vitro* data? (ii) how does the discriminating potential of superferns quantitatively compare with the fern features? (iii) how diverse is the classifier ensemble created by our proposed approach?

**Experimental Setup:** For each detected spots  $\hat{\mathbf{s}}$ , we claim a true detection if there is a ground truth spot within a small distance, i.e.,  $\|\hat{\mathbf{s}} - \mathbf{s}_i\| < 0.5$  pixel. The ROC, and Area Under the Curve (AUC) are the evaluation metrics. For the



**Fig. 6.** Detection performance comparisons and with various components.

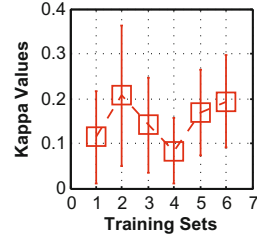
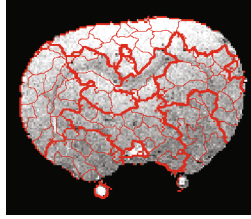
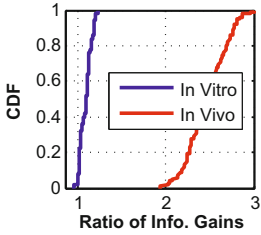


**Fig. 7.** Spot detection examples: (a) true detection, (b) false negative, and (c) false positive of *in vivo* data. Each column shows two consecutive slices of a spot. Note the amount of appearance variations in *in vivo* data. (d) visualization for a 3 sliced spot.

5-scans *in vivo* data, we adopt a leave-two-out scheme such that our testing set contains one labeled and one spotless scan. This creates six pairs of training and testing sets, which allows us to compute the error bar of ROC. For the 4-scans *in vitro* data, three pairs of training and testing sets are formed such that the naïve scan remains in the testing set accompanied by every other scan once. Since [4] and [9] are the most relevant examples of MRI cell detection using learning-based and rule-based methods, we implement them as the baselines. We experimentally determine  $\tau = 9$ ,  $h = 215$ ,  $N_0 \in [6, 9]$  for *in vivo* data, and  $U \in [200, 2000]$  and  $\tilde{U} \in [20, 60]$  depending on the size of brain regions.

**Performance and Comparison:** As shown in Fig. 6 (a,b), the proposed method outperforms two baselines with an average AUC of 98.9% (*in vitro*) and 89.1% (*in vivo*). The improvement margin is especially larger at lower FPRs, which are the main operation points in practice. Fig. 6 (c) shows that with *in vivo* data, by using ferns instead of superferrns or by making no partitions of the brain region, we observe a decrease in AUC to 87.1% and 85.3%, respectively. We also applied the standard implementation of C4.5 decision tree to our data, since decision trees naturally splits data based on features, however, the obtained results were significantly inferior to our approach. Fig. 7 shows three types of spot detection results with our method. The appearance and shape variations among the spots clearly show the challenge of this problem.

**Superferrns vs. Ferns:** To further illustrate the strength of the novel superferrns feature, we compare the discriminating potential of superferrns with ferns, *regardless* of the classifier design. Information gain (IG) is a standard tool to measure the worth of a feature, where a higher IG indicates its higher discrimi-



**Fig. 8.** The ratio of IGs. **Fig. 9.**  $\tilde{V}_{z,\cdot}$  & 8 partitions. **Fig. 10.** Classifier diversity.

nating potential. Given a set of 50 randomly generated offsets  $\{\mathbf{o}_i\}$ , we calculate their superferns features on the *in vitro* data including both spots and non-spots, which allows us to compute  $A_s$  containing the IG of each superferm. The *same* offsets are applied to the fern features and then their corresponding set of IGs  $A_f$  is computed. Finally we compute the ratio of both IGs,  $\frac{A_s(i)}{A_f(i)}$ , for each offset  $\mathbf{o}_i$ , and show the Cumulative Density Function (CDF) of 50 ratios in Fig. 8. Using 100 random offsets, the same experiment is repeated for the *in vivo* data. The fact that almost all ratios are larger than 1 shows the superiority of superferns.

**Diversity Analysis:** Our classification framework includes an ensemble of classifiers, one for each partition. Fig. 9 shows for one slice the  $N_0 (= 8)$  partitions and the enclosed large superpixels  $\tilde{V}_{z,\cdot}$ . Since diverse discriminative features are likely to be utilized in different partitions, learning on disjoint partitions should favor high diversities among classifiers – an indicator for effective classification. To evaluate the diversity of our classifier ensemble, we use the standard Cohen’s kappa value [6], which ranges from 0 to 1, with a lower value indicating a higher diversity. For each of six *in vivo* training sets, we compute  $\frac{N_0(N_0-1)}{2}$  kappa values, each between a pair of classifiers learned on different partitions. Fig. 10 shows their mean and standard deviation for each training set. According to [6], our kappa values is very low, indicating the high diversity of learned ensemble.

## 5 Conclusions and Future Work

This paper presents the first comprehensive study on learning-based cell/spot detection in MRI. We propose a novel framework that employs a superpixel-based 3D spot model, extracts *superferns* features, and utilizes a partition-based classifier ensemble. We collected a total of nine MRI scans with more than 16 thousand labels for evaluation. Experimental results demonstrate the superiority of our approach. In future, we intend to investigate two directions: (i) further improving the detection accuracy by learning more contextual information of the brain; (ii) extend the spot detection from brain to other organs, such as liver.



## References

1. Bouckaert, R., Frank, E., Hall, M., Kirkby, R., Reutemann, P., Seewald, A., Scuse, D.: Weka manual for version 3-7-8 (2013)
2. Heyn, C., Ronald, J.A., Mackenzie, L.T., MacDonald, I.C., Chambers, A.F., Rutt, B.K., Foster, P.J.: In vivo magnetic resonance imaging of single cells in mouse brain with optical validation. *Magnetic Resonance in Medicine* 55(1), 23–29 (2006)
3. Liu, M., Tuzel, O., Ramalingam, S., Chellappa, R.: Entropy rate superpixel segmentation. In: *IEEE CVPR*, pp. 2097–2104 (2011)
4. Mori, Y., Chen, T., Fujisawa, T., Kobashi, S., Ohno, K., Yoshida, S., Tago, Y., Komai, Y., Hata, Y., Yoshioka, Y.: From cartoon to real time MRI: in vivo monitoring of phagocyte migration in mouse brain. *Scientific Reports* 4 (2014)
5. Ozuysal, M., Calonder, M., Lepetit, V., Fua, P.: Fast keypoint recognition using random ferns. *IEEE T-PAMI* 32(3), 448–461 (2010)
6. Rodriguez, J.J., Kuncheva, L.I., Alonso, C.J.: Rotation forest: A new classifier ensemble method. *IEEE T-PAMI* 28(10), 1619–1630 (2006)
7. Shapiro, E.M., Sharer, K., Skrtic, S., Koretsky, A.P.: In vivo detection of single cells by MRI. *Magnetic Resonance in Medicine* 55(2), 242–249 (2006)
8. Slotkin, J.R., Cahill, K.S., Tharin, S.A., Shapiro, E.M.: Cellular magnetic resonance imaging: nanometer and micrometer size particles for noninvasive cell localization. *Neurotherapeutics* 4(3), 428–433 (2007)
9. Smal, I., Loog, M., Niessen, W., Meijering, E.: Quantitative comparison of spot detection methods in fluorescence microscopy. *IEEE T-MI* 29(2), 282–301 (2010)